Can we predict good drugs?

the application of
~ big database to R&D of drugs ~


10th November 2012 (Sat.)
Takuya Matsumoto
Combined FDA-approved new molecular entities (NMEs) versus R&D spending for the top nine largest pharmaceutical companies.

**NME**: a drug that contains no active moiety that has been approved by the FDA in any other application submitted. NME includes biologicals and vaccines.  

NMEs success rate by phase. Combined R&D survival by development phase for the top fourteen largest pharmaceutical. Approval data is based on approval of NME by a regulatory authority in a major market (EU, US or Japan).

Next to lack of efficacy, **adverse drug reactions (ADRs)** are the leading cause for attrition in clinical trials of new drugs.

Some ADRs are caused by modulation of the primary target of a drug, others result from non-specific interactions of reactive metabolites. In many cases, however, ADRs are caused by **unintended activity at off-targets**.

Cyclooxygenase (COX) has two well-studied isoforms, called COX-1 and COX-2. COX-1 mediates the synthesis of prostaglandins responsible for protection of the stomach lining, while COX-2 mediates the synthesis of prostaglandins responsible for pain and inflammation. By creating "selective" NSAIDs that inhibit COX-2, but not COX-1, the same pain relief as traditional NSAIDs is offered, but with greatly reduced risk of fatal or debilitating peptic ulcers.
Today’s topic

Predicting unintended off-target toxicity by informatics methods!

Informatics: the science of information
It studies the representation, processing, and communication of information in natural and artificial systems.

Bioinformatics (DNA/RNA, protein, etc.)
Chemoinformatics (ligand, drugs)
What's SEA?

SEA = the Similarity Ensemble Approach

1. Score the similarity of all ligands pairs between two proteins with Tanimoto efficient (Tc) (0 (complete dissimilarity) < Tc < 1 (identity))
2. Sum up all Tc score (the raw similarity score)
3. Calculate expectation values and evaluate the similarity of two proteins

= 0.2 + 2 × 0.3 + 0.5 = 1.3 (the raw similarity score)

the total similarity score between protein A (ligand α & β) & B (ligand 1 & 2)

the total number of
Protein: 246
Ligand: 65,241
(Pair: 5.07 x 10⁹ ÷ 65,241²)
**Tanimoto coefficient (Tc)**

**Basic idea**

\[ T_s(X, Y) = \frac{\sum_i (X_i \land Y_i)}{\sum_i (X_i \lor Y_i)} \]

**Examples**

Tylenol (paracetamol) \[ \frac{8 \text{ in common}}{18 \text{ total}} = 0.44 \]

Advil (ibuprofen)

\[ \begin{align*}
\text{Chlorotrianisene} & \quad 0.31 \quad \text{COX-1} \\
\text{Indomethacin} & \quad \text{Medrysone} \\
& \quad \text{CHEMBL307710}
\end{align*} \]
The enzymes thymidylate synthase (TS) and dihydrofolate reductase (DHFR) both recognize folic acid derivatives and are inhibited by antifolate drugs. Despite this, two enzymes have no substantial sequence identity and are structurally unrelated.

Thrombin: serine protease protein (unrelated to DHFR)

- For most ligand pairs the Tc was low, in the 0.2 to 0.3 range (insubstantial similarity). This was true even when comparing a set to itself. DHFR versus DHFR (red), 80.4% (0.1 to 0.4 range), 4.7% (0.6–1.0 range) and 0.5% (1.0)

- The raw similarity score: the sum of ligand pair Tcs over all pairs with Tc ≥ 0.57

*Nature Biotechnology* 2007, 2, 197.
Quantifying SEA

\[ P = 1 - \exp(Kmne^{-\chi}) \text{ where } \chi \text{ is similarity score} \]

\[ \chi \text{ is the raw similarity score (?)} \]

The smaller, the more similar

Folic acid

On average, any given receptor was similar to only 5.8 other receptors with an expectation value $<10^{-10}$.

Clusters of biologically related targets may be observed, as no explicit biological information, only ligand information, is used to calculate the cross-target similarity.

*Nature Biotechnology* 2007, 2, 197.
Comparison of sequence and ligand-based protein similarity

Red: pairs with strong ligand-set similarity but weaker sequence similarity.

Dark gray: pairs with strong sequence similarity but comparatively lower ligand-set similarity

White: pairs where pharmacological and sequence similarity approaches agree (either positively or negatively)

(a) overall difference heat map
(b) folate-recognition enzymes and adenosine-binding enzymes
(c) GPCRs, ion channels and nuclear hormone receptors

*Nature Biotechnology* 2007, 2, 197.
Predicting and testing drug promiscuity

Methadone is known to have dual specificity for μ-opioid receptors and NMDA.

<table>
<thead>
<tr>
<th>Query</th>
<th>Rank</th>
<th>Activity class</th>
<th>E-value</th>
<th>Max Tc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methadone&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1</td>
<td>Antimuscarinic</td>
<td>4.45×10⁻⁵⁰</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Muscarinic M3 antagonist</td>
<td>1.22×10⁻¹¹</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Opioid agonist</td>
<td>1.84</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>NMDA receptor antagonist</td>
<td>9.04</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Muscarinic (M1) agonist</td>
<td>61.9</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Cyclooxygenase inhibitor</td>
<td>12.1</td>
<td>0.61</td>
</tr>
</tbody>
</table>

Antagonism of M3 muscarinic receptors by the μ-opioid agonist methadone in a direct binding assay

Antagonism of M3 muscarinic receptors by functional assay

*Nature Biothechnology* 2007, 2, 197.
Small Summary 1

SEA = the Similarity Ensemble Approach

Protein A

ligand α

0.5

ligand β

0.3

0.3

0.2

Protein B

ligand 1

ligand 2

DHFR set similarity distributions

Tylenol
(paracetamol)

8 in common
18 total = 0.44

Advil
(ibuprofen)
Predicting new molecular targets for known drugs

**Drugs**

\[
878 + 2,787 = 3,665
\]

- **FDA-approved drugs**: 878
- **investigational drugs**: 2,787
- **total**: 3,665

**Data base**

same as the previous study (*Nature Biothechnology 2007*, 2, 197.)

- **246 proteins**
  - (65,241 ligands)

**Total**

- **901,590 protein-drug pairs** (3,665 x 246)

**Result**

- **901,590 protein-drug pairs**
  - **6,928 pairs** with \( E \)-values better than \( 1 \times 10^{-10} \)
  - **3,832 pairs** unknown, biologically interesting
  - **184 pairs**
    - **40 pairs** known
    - **30 pairs** tested
  - **23 pairs** \( K_i < 15 \ \mu M \)

New targets as primary sites of action

<table>
<thead>
<tr>
<th>Drug</th>
<th>Pharmacological action</th>
<th>$E$-value</th>
<th>Predicted target</th>
<th>$K_i$ (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="" alt="Chemical Structure" /></td>
<td>Antihistamine (H$_1$ receptor)</td>
<td>$5.7 \times 10^{-67}$</td>
<td>5-HT$_{5A}$ antagonist (serotonergic receptor)</td>
<td>130</td>
</tr>
</tbody>
</table>

- Used since the 1950s as an antihistamine, Fabahistin is now being investigated for Alzheimer’s disease.
- **Fabahistin binds predicted new, off-target(5-HT$_{5A}$ receptor) much stronger than its canonical H$_1$ receptor target.**
- Its activity against 5-HT$_{5A}$ and related serotonergic receptors may have implications for Fabahistin’s role as an Alzheimer’s disease therapeutic.

**Off-targets as side-effect mediators**

<table>
<thead>
<tr>
<th>Drug</th>
<th>Pharmacological action</th>
<th>$E$-value</th>
<th>Predicted target</th>
<th>$K_I$ (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Motilium structure" /></td>
<td>Antiemetic; peristaltic stimulant (dopamine D$_2$ receptor)</td>
<td>$5.48 \times 10^{-11}$</td>
<td>$\alpha_1$ adrenergic blocker</td>
<td>$\alpha_{1A}: 71$; $\alpha_{1B}: 530$; $\alpha_{1D}: 710$</td>
</tr>
</tbody>
</table>

- Motilium achieves peak plasma concentrations ($C_{max}$) of 2.8 μM after intravenous administration.

- This formulation was withdrawn owing to adverse cardiovascular effects, with the FDA citing cardiac arrest, arrhythmias and sudden death.

- Although Motilium binds the hERG potassium ion channel with a half-maximum inhibitory concentration ($IC_{50}$) of 5 μM, the 71–710nM affinities observed here against $\alpha_{1A}$, $\alpha_{1B}$ and $\alpha_{1D}$ may also contribute to these cardiovascular effects.

*Nature* **2009, 462, 175.**
**Drug binding across major protein boundaries**

<table>
<thead>
<tr>
<th>Drug</th>
<th>Canonical target</th>
<th>E-value</th>
<th>Predicted target</th>
<th>$K_i$ (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xenazine</td>
<td>VMAT2 (transporter)</td>
<td>$1.4 \times 10^{-61}$</td>
<td>$\alpha_2$ adrenergic receptor (GPCR)</td>
<td>$\alpha_{2A}, 960; \alpha_{2C}, 1.3 \times 10^3$</td>
</tr>
<tr>
<td>Rescriptor</td>
<td>HIV-1 reverse transcriptase (enzyme)</td>
<td>$1.1 \times 10^{-20}$</td>
<td>Histamine $H_4$ receptor (GPCR)</td>
<td>$5.3 \times 10^3$</td>
</tr>
<tr>
<td>Vadilax</td>
<td>NMDAR (ion channel)</td>
<td>$5.1 \times 10^{-32}$</td>
<td>$\mu$-opioid receptor (GPCR)</td>
<td>$1.4 \times 10^2$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$2.0 \times 10^{-4}$</td>
<td>5-HTT; serotonin transporter (transporter)</td>
<td>77</td>
</tr>
<tr>
<td>RO-25-6981</td>
<td>NMDAR (ion channel)</td>
<td>$1.5 \times 10^{-6}$</td>
<td>5-HTT; serotonin transporter (transporter)</td>
<td>$1.4 \times 10^3$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$1.9 \times 10^{-6}$</td>
<td>Dopamine $D_4$ receptor (GPCR)</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$3.6 \times 10^{-6}$</td>
<td>NET, noradrenaline transporter (transporter)</td>
<td>$1.3 \times 10^3$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$9.1 \times 10^{-5}$</td>
<td>$\kappa$-opioid receptor (GPCR)</td>
<td>$3.1 \times 10^3$</td>
</tr>
</tbody>
</table>

- The protein target with the highest sequence similarity to any of a drug’s known targets is rarely predicted by the SEA approach.
- Rather, the target predicted by ligand similarity is typically well down in the sequence-similarity ranking.

3,665 (FDA)-approved and investigational drugs were compared against 246 targets. 364 additional off-targets for 158 drugs are predicted with $E$-values better than $1 \times 10^{-50}$, whereas 1,853 new off-targets are predicted with $E$-values better than $1 \times 10^{-10}$. This compares to the only 972 off-target activities already annotated in the databases. Prediction includes some interesting new off-targets such as:

(1) the new targets contribute to the primary activity of the drug
(2) the new targets may mediate drug side effects
(3) the new targets are unrelated by sequence, structure and function to the canonical targets.

Large-scale prediction and testing of drug activity on side-effect targets

1. Calculate $E$-value by SEA methods, predict new drug-off-target and confirm by in vitro experiment (similar as the previous study (Nature 2009, 462, 175.))

2. Quantify the relationships between protein targets and adverse drug reactions (ADRs) by the use of enrichment factor (EF) (different point from the previous study)

3. Create a drug–target–ADR network

Large-scale prediction and testing of drug activity on side-effect targets

Drugs & Targets

656 FDA-approved drugs listed in ChEMBL

\[ \times \quad 73 \quad = \quad 47,888 \quad \text{total} \]

with established association of ADRs, for which assays were available at Novartis

Sea

1,644 pairs with \( E \)-values better than \( 1 \times 10^{-4} \)

403 pairs listed in ChEMBL

1,241 pairs

348 pairs listed in other database

694 pairs tested at Novartis

893 pairs

New drug-off-target predictions

<table>
<thead>
<tr>
<th>Drug</th>
<th>Closest chEMBL molecule</th>
<th>Tc value</th>
<th>Target</th>
<th>SEA E value</th>
<th>IC₅₀ (µM)</th>
<th>Closest known target</th>
<th>BLAST E value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alosetron</td>
<td></td>
<td>0.25</td>
<td>HTR2B</td>
<td>10.5 × 10⁻¹⁷</td>
<td>0.02</td>
<td>KCNH7</td>
<td>3.6 × 10⁻²</td>
</tr>
<tr>
<td>Aprindine</td>
<td></td>
<td>0.38</td>
<td>HRH1</td>
<td>5.0 × 10⁻²⁵</td>
<td>0.78</td>
<td>SCN5A</td>
<td>3.3 × 10⁻¹</td>
</tr>
<tr>
<td>Chloroethanol</td>
<td></td>
<td>0.36</td>
<td>COX-1</td>
<td>1.9 × 10⁻¹⁷</td>
<td>0.16</td>
<td>ESR1</td>
<td>9.0 × 10²</td>
</tr>
<tr>
<td>Cemastine</td>
<td></td>
<td>0.31</td>
<td>SLC6A4</td>
<td>1.1 × 10⁻¹⁴</td>
<td>0.42</td>
<td>KCNH2</td>
<td>6.1 × 10¹</td>
</tr>
<tr>
<td>Dyclonine</td>
<td></td>
<td>0.37</td>
<td>DRD4</td>
<td>1.5 × 10⁻¹⁷</td>
<td>4.1</td>
<td>SLC6A3</td>
<td>2.3 × 10²</td>
</tr>
<tr>
<td>Fasudil</td>
<td></td>
<td>0.31</td>
<td>ADRA2A</td>
<td>1.1 × 10⁻⁷</td>
<td>4.0</td>
<td>CCR2</td>
<td>1.5 × 10⁻⁹</td>
</tr>
<tr>
<td>Prenylamine</td>
<td></td>
<td>0.30</td>
<td>OPRM1</td>
<td>3.2 × 10⁻⁶⁵</td>
<td>7.9</td>
<td>SCN5A</td>
<td>3.3 × 10⁻¹</td>
</tr>
</tbody>
</table>

SEA or 1NN

1NN = one-nearest neighbor model

Adjusted hit rate = (number of true positives + 1) / (number of total predictions + 1)

(number of total prediction = number of true positive + number of false positive)

Are new off-targets predictable?

Of the 151 (ref. slide 21) new off-target predictions, 39 (26%) had BLAST E-values greater (worse) than $10^{-5}$, suggesting the previously known targets shared no sequence similarity with the new off-targets.

**Associating in vitro targets with ADRs**

SEA relationships between drugs and targets to assess the potential clinical relevance of the discovered targets of drugs systematically...

A quantitative score that associated *in vitro* activity with patient ADRs (a score between targets and ADRs)

**Solution:** Enrichment Factor (EF)

$$EF = \frac{p}{(A \times T \div P)}$$

in which *p* is the co-occurrence of target X and ADR Y, *A* is the number of times ADR Y was linked to any drug–target pair, *T* is the number of times target X was linked with any drug–ADR pair, and *P* is the total number of target–ADR pairs.

45 drugs (*p*) which have the ADR epigastralgia and interact with COX-1

6,046 drug–target pairs (*A*) where the ADR epigastralgia r was linked with

2,188 drug–ADR pairs (*T*) where COX-1 was linked with

681,797 target–ADR pairs overall (*P*)

Thus the pair epigastralgia–COX-1 was enriched **2.3-fold above random**

*Nature 2012, 486, 361.*
Of the 151 confirmed (ref. slide 21) new drug–target associations tested at Novartis, 82 were significantly associated with one or more ADR, resulting in a total of 247 drug–target–ADR links.

In 116 cases, the enrichment factor (EF) of the new drug–target–ADR link was stronger than that for any previously known target.
Drug–Target–ADR network

Grey: known targets
Blue: newly predicted targets
Orange, Red: ADRs associated with each targets (Red: ADRs are significantly (EF>1) associated with the new off-targets.) (Targets related by sequence are connected by grey edges.)

Demonstration of an association in an accepted in vivo biomarker

This was the first example that a synthetic steroid acted on COX-1 enzyme!
Target promiscuity

Drug promiscuity

Summary

SEA = the Similarity Ensemble Approach
EF = Enrichment Factor

Drug-Target-ADR network

Comments

- Only some side effects fall into the remit of this approach, which assumes an off-target mechanism.

- Almost 46% of the predicted drug–target associations were disproved, but they were just as often confirmed by experimental ways.

- The method was used automatically at scale, without human intervention. Pragmatically, the ability to calculate drug–target–ADR networks provides a tool to anticipate liabilities among candidate drugs being advanced towards the clinic, or yet earlier, for prioritization of chemotypes in preclinical series.

- The use of Big Data will be dramatically accelerated in almost all fields.
How to measure “Drug-Likeness”? A new measure taking the place of ‘the Lipinski’s Rule of Five’

Oral drug & Lipinski's rule of five

Oral drug is the best, thus the most important way to dose drugs.

empirical criteria whether a small organic molecule is suitable for a oral drug

Lipinski's rule of five

Lipinski’s rule of five

- Its molecular weight is less than 500.
- The compound’s lipophilicity, expressed as a quantity known as \(\log P\) (the logarithm of the partition coefficient between water and 1-octanol), is less than 5.
- The number of groups in the molecule that can donate hydrogen atoms to hydrogen bonds (usually the sum of hydroxyl and amine groups in a drug molecule) is less than 5.
- The number of groups that can accept hydrogen atoms to form hydrogen bonds (estimated by the sum of oxygen and nitrogen atoms) is less than 10.

Nature 2012, 481, 455.
The Implementation of Rules

The rules are only predictive of oral bioavailability (the absorption by passive diffusion of compounds through cell membranes). Due to their simplicity, the rules are widely used by medicinal chemists to predict not only the absorption of compounds, as Lipinski originally intended, but also overall drug-likeness. Despite Lipinski’s recommendation that the rule be considered as a guideline, in reality it is used routinely to filter libraries of compounds. The implementation of rules as filters means that no discrimination is achieved beyond a qualitative pass or fail—all compounds that comply with the rules are considered equal, as are all that breach them.

Nature 2012, 481, 455.
Quantifying drug-likeness

To quantify compound quality, the concept of desirability was applied to provide a quantitative metric for assessing drug-likeness, which we call the quantitative estimate of drug-likeness (QED). QED values can range from 0 (all properties unfavorable) to 1 (all properties favorable).

Desirability takes multiple numerical or categorical parameters measured on different scales and describes each by an individual desirability function. These are then integrated into a single dimensionless score. In the case of compounds, a series of desirability functions $d$ are derived, each of which corresponds to a different molecular descriptor. Combining the individual desirability functions into the QED is achieved by taking the geometric mean of the individual functions, as shown in equation.

$$QED = \exp\left(\frac{1}{n} \sum_{i=1}^{n} \ln d_i\right)$$

**Weighted QED**

$$QED_w = \exp\left(\frac{\sum_{i=1}^{n} w_i \ln d_i}{\sum_{i=1}^{n} w_i}\right)$$

**Asymmetric Double Sigmoidal (ADS) functions**

$$d(x) = a \left[\frac{b}{1 + \exp\left(-\frac{x-c+d}{2}\right)}\right] \left[1 - \frac{1}{1 + \exp\left(-\frac{x-c-d}{2}\right)}\right]$$

(a - f: constant values)

Mr: molecular weight

Histograms of selected properties

Histograms of eight selected molecular properties for a set of 771 orally absorbed small molecule drugs. molecular weight Mr (a), lipophilicity estimated by atom-based prediction of ALOGP (b), number of HBDs (c), number of HBAs (d), PSA (e), number of ROTBs (f), number of AROMs (g) and number of ALERTS (h).
The Lipinski-compliant areas are shown in pale blue in (a), (b), (c) and (d).

Prediction of the drug-likeness of proteins' ligand

Not all ligand-binding sites have the appropriate physicochemical and topological properties to bind small-molecule drugs non-covalently with sufficient affinity. Binding sites that do have these characteristics are described as druggable (this definition is independent of any wider biological considerations). QED provides an efficient means to quantify and rank the druggability of targets according to the chemical attractiveness of their associated ligands. In other words, proteins whose ligands had the highest QED scores should be the most chemically tractable targets for drug discovery, because their known ligands are the most drug-like.

Structural diversity networks

(a) a target for which the associated bioactive compounds are neither drug-like nor diverse
(b) a target for which the associated bioactive compounds are diverse, but not drug-like
(c) a target for which the associated bioactive compounds are drug-like, but not diverse
(d) a target for which the associated bioactive compounds are both drug-like and diverse

In each of the networks compounds are represented as nodes and are coloured by their respective QED values. An edge connects nodes if they are structurally similar (defined by a Tanimoto coefficient ≥ 0.7).
Top human targets by three different ranking schemes

<table>
<thead>
<tr>
<th>Target (UniProt)</th>
<th>Mean QED</th>
<th>Mean QED best cluster</th>
<th>Proportion clusters with mean QED &gt; 0.796</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free fatty acid receptor 2 (O15552)</td>
<td>0.861</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sodium channel (Q9NY72, Q60939, Q8WT1, Q07699)</td>
<td>0.849</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voltage-gated potassium channel (P15382, P51787)</td>
<td>0.833</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phosphodiesterase 9A (Q76083)</td>
<td>0.820</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aldo-keto-reductase family 1 member C3 (P42330)</td>
<td>0.812</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cholinergic receptor, nicotinic, beta 1 (muscle) (Q81246)</td>
<td>0.811</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sorbitol dehydrogenase (Q00796)</td>
<td>0.850</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sodium channel protein type IV alpha subunit (P35499)</td>
<td>0.809</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endothelial lipase (Q9YSX9)</td>
<td>0.804</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vesicular acetylcholine transporter (Q16572)</td>
<td>0.798</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuropeptide Y receptor type 5 (Q15761)</td>
<td>0.935</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serotonin transporter (P31645)</td>
<td>0.932</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serotonin 1a (5-HT1a) receptor (P09808)</td>
<td>0.932</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noradrenaline transporter (P23975)</td>
<td>0.932</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dopamine transporter (P07959)</td>
<td>0.931</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Histamine H1 receptor (P35367)</td>
<td>0.928</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dopamine D3 receptor (P35462)</td>
<td>0.927</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dopamine D4 receptor (P29197)</td>
<td>0.925</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thromboxane-A synthase (P24557)</td>
<td>0.921</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serotonin 2c (5-HT2c) receptor (P28335)</td>
<td>0.917</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vesicular acetylcholine transporter (Q16572)</td>
<td>0.714</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phosphodiesterase 7A (R3946)</td>
<td>0.667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Melatonin receptor 1A (P34813)</td>
<td>0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Melatonin receptor 1B (P49298)</td>
<td>0.462</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noradrenaline transporter (P23975)</td>
<td>0.453</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Histamine H4 receptor (Q93H4)</td>
<td>0.444</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dopamine transporter (Q09959)</td>
<td>0.409</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serotonin 7 (5-HT7) receptor (P46969)</td>
<td>0.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuronal acetylcholine transporter protein beta-4 subunit (P11092)</td>
<td>0.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuronal acetylcholine transporter protein alpha-7 subunit (P36544)</td>
<td>0.4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Left  | ranking targets by the mean QED of their associated ligands

Center | ranking targets by the mean of the most drug-like active series (clusters)

Right  | ranking targets by the degree of enrichment of drug-like series (type (d) target in the previous slide) (targets are ranked by the proportion of active series that have a mean QED above that of the top 10% of the ChEMBL database (0.796)).

The mean QED for all targets in the list is 0.478. For the targets of approved drugs the mean QED is 0.492 and for the targets of approved oral drugs the mean QED is 0.539. Drug targets are, indeed, enriched towards the more highly desirable targets, with 70% of the drug targets found in the top 50% of the prioritized target list.